

DOCUMENT RESUME

ED 458 297

TM 033 473

AUTHOR Chartrand, Anne
TITLE Whose Standards Are They, Anyway?
PUB DATE 2000-00-00
NOTE 9p.
PUB TYPE Opinion Papers (120)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Achievement; *Academic Standards; *Accountability; Definitions; Educational Assessment; *Educational Change; Educational Quality; Elementary Secondary Education
IDENTIFIERS *Reform Efforts

ABSTRACT

Over the years, the definitions of educational standards have become more varied, and the issue of what standards really mean has become more confusing. The answer to the question, "For whom are content standards developed?" helps determine who takes them seriously. Establishing clear content standards plays a critical role in an instructional program, and any discussion about establishing content standards must include a discussion of how to assess standards. The specifications for an assessment should come from the content standards if the state or school district truly believes the content standards are essential skills and knowledge students should learn. If the emphasis is all on a norm-referenced test, it is at best difficult, if not impossible, to report exactly what the student knows and can do. The overemphasis on a single test score remains. In many cases the emphasis on increasing test scores comes at the expense of the best educational practices for all students. In considering the demand for increased test scores, one must wonder when they will be high enough. Should all schools be expected to have the same amount of growth or the same score levels? On the positive side, large scale assessments are developed more carefully than they were years ago. The underlying question of whether the emphasis on standards and assessment is having a positive effect on instructional improvement and achievement overall is being examined, and the creation of this dialogue may be the best result of the efforts toward higher standards. Neither identifying standards nor administering assessments will make students "smarter," but discussing the roles of standards and assessments can result in the improvement of instruction. (SLD)

Whose Standards are They, Anyway?

Anne Chartrand
National Computer Systems

TM0333473

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

✓ This document has been reproduced as
received from the person or organization
originating it.

☐ Minor changes have been made to
improve reproduction quality.

• Points of view or opinions stated in this
document do not necessarily represent
official OERI position or policy.

PERMISSION TO REPRODUCE AND
DISSEMINATE THIS MATERIAL HAS
BEEN GRANTED BY

A. Chartrand

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

1

WHOSE STANDARDS ARE THEY, ANYWAY?

Anne Chartrand, Ed. D.
National Computer Systems

Over the years, terminology and definitions for what we currently are calling standards have been tweaked, redefined, adjusted to new circumstances, and combined, often to the point of confusion. Many years ago, when discussing "setting standards," we clearly were discussing the pros and cons of cut score methodologies suggested by researchers such as Ebel, Nedelsky, and the many ways to modify Angoff. Today, standard setting has come to be associated more with the articulation of curriculum, as well as indicating performance levels established on assessment results. The definitions of standards have become more varied and appears to change based on situation, purpose, objective, and interpreter. For example, just examine some of the past years. We could begin with behavioral objectives, and move on to goals, objectives, scope and sequence, competencies (minimum and otherwise), essential skills (as opposed to nonessential skills), curriculum frameworks, courses of study ... you know, standards. Clarification also would be helpful for the variety of proficiency levels used across the country. Can we adequately describe a student who is basic, proficient, above average, below average, advanced, level 1, level 4, below basic ... you know, standards.

This is not meant to be glib, but if we continue to confuse the issue, how can we expect the public to understand what we are trying to convey? Educators and policymakers constantly are making the statement that they want to "raise the standards." Does this mean the skills and knowledge expected of students are to be more difficult, or are we demanding an increase in test scores, a higher "cut" score on the same standards, or an increase in the number of students in higher proficiency levels? How do we best explain "standards of performance"?

For whom are the content standards developed? The answer to this question helps determine who takes them seriously. If the standards truly are a guide for teachers to use in everyday instruction, then the standards become the curriculum and those in individual schools must pay attention. If the standards are simply a list of skills contained in a document to prove to policymakers that standards exist, then they often are meaningless to teachers. Although Gandal reports that there has been improvement in the revising of state standards that previously were identified as vague and weak, his report still concludes that the efforts were "far from acceptable" (Gandal, 1996). In his recommended guidelines for the large-scale assessment community, Popham opines that most of the state-level content standards that he has seen "represent little more than pious wish-lists at generality levels little better than the gunky state-level curricular syllabi of yesteryear" (Popham, 1999). If standards are vague, nondescriptive, and not easily converted to instructional activities, then they are useless to the teachers.

Parents and students often believe the textbooks to be the standards because tests and grades are based so heavily upon textbooks. Parents and students also may have the impression that teachers are establishing their own content standards as well as the standards of acceptance, i.e., grades. E.D. Hirsch published what he believes to be the "core knowledge" that all students should learn by grade level. His contention is that, while most curriculum documents are vague,

there is some essential common content that is specific and that should be available to all students regardless of where they live. He believes that this core knowledge should be about 50% of the curriculum in schools (O'Neil, 1999). Not only are there many schools that have adopted this concept, but also some military parents who use it adamantly to determine whether they believe their children are receiving an adequate education as the family moves from location to location. Their determination of essential standards is derived, not from the state or school documents, but from what is happening in the current classroom and how it compares with their experiences in other towns.

Establishing clear content standards certainly plays a critical role in an instructional program. The value of content standards can vary, however, based upon who developed them, for whom they are intended, and the realism of their application in the classroom. Some of the common problems that must be avoided are related to identifying standards for students that are too broad or too narrow, too rigorous or too trivial, too many or too few. If standards are merely the skills measured on a statewide test, then it is evident that the test controls the educational system. Standards must be communicated to and understood by both the teachers and the students. They must be able to be successfully taught, contribute to thinking and reasoning, broad enough to be sampled but identified sufficiently well enough to be taught (Schomaker and Marzano, 1999), and given the attention to allow teachers to be proficient in their teaching. This is not an easy task. We can give so much attention to the development and refinement of standards that we neglect the pedagogical support system that must impart them.

Any discussion about establishing content standards also must include a discussion of how to assess standards. Some of the same issues encountered in developing content standards also apply in choosing or designing a measure for the standards. The specifications for an assessment should come directly from the content standards if the state or district truly believes the content standards are the essential skills and knowledge students should learn. The assessments, like the content standards, should not be too broad or too narrow (item-focused).

Even if a state or district has acclaimed content standards, if all emphasis is on a norm-referenced test (NRT), it is at best difficult, if not impossible, to report exactly what the student knows and can do, but rather how that student compares in performance with other students in a norm group. In Popham's continuing discussion of the shortcomings of NRTs, he points out that in the "quest for score variance in a standardized achievement test, items on which students perform well are often excluded. However, items on which students perform well often cover the content that, because of its importance, teachers stress" (Popham, 1999, p. 12). There are other problems when an NRT is taking the major media role in large-scale assessment. Regardless of what standards have been established and, in some cases, what instruments measure them, the magic of the 50th percentile gets the most attention. Politicians (and many others) still equate increased test scores with increased student knowledge. Worthiness is defined by the national average, and the demise or salvation of a school can be determined by a single score interpretation. In one particular instance, a method was established to identify those schools "in danger." In order to be "cleared," a certain number of students had to move out of a specific, narrowly defined, band of percentile ranks. It was quite possible for the school to improve significantly its overall test scores, yet remain on the danger list because of this one small group of students. The achievement of the other students was ignored. When only a few

students' scores can "make or break" a school's rating, the philosophical practice of teaching *all* students becomes less of a reality.

The overemphasis on a single test score still prevails. Many states have developed customized assessments that are in congruence with their standards and have followed intense development procedures to assure their assessment's reliability and validity. When the scores are released, however, attention befalls the low-scoring schools, where teaching to the test is then predictable. Although many educators would like to deny that this is the case, states where either their state-developed tests or NRTs are the sole determiner of the quality of a school, teachers readily will admit that all they do is concentrate primarily on what is measured on the test. Other activities and subject areas are ignored or dropped to the bottom of a priority list. Accountability, rather than instructional prowess, has control in many places. Have we created slaves to increase test scores at the expense of the best educational practices for all students? Yes, and this is certainly not a new problem. More effort needs to be expended in studying the few locations that appear not to have fallen victim to this undesired outcome so that we can determine the feasibility of assisting others who are overwhelmed with this predicament.

In considering the demand for increased test scores, when will they be high enough? This is not to suggest at all that we are not in need of vast improvement in many areas, but is it possible that the continuing cry for increased test scores may result in demanding cognitive skills that are unrealistic for many students? David Hoff addresses the issue of the predictability of ups and downs in test scores in a recent article. He cites several sources and reasons for test scores to "start low, rise quickly for a couple of years, level off for a few more, and then gradually drop over time" (Hoff, 2000). He goes on to quote researchers who agree that schools take care of the easy things first and neglect to make systemic changes that can make a difference in the long run. The cycle continues when policymakers then get anxious and search for a new test. How much growth is reasonable? How much higher can the standards be without becoming illogical to those who teach and to those who must demonstrate knowledge? Should all schools be expected to have the same amount of growth or the same score levels?

On a positive side, large-scale assessments, for the most part, are developed much more carefully than many years ago, and with much more input from content and instructional leaders who take their tasks quite seriously. Closer attention is paid to bias issues, and procedures and consensus attempts are put into place that involve many more people than just a test developer. Piloting, field testing, and scrutiny of every item for content are intense, many times due to the fear of litigation, but add to the quality of the instrument. The addition of performance assessments has contributed a component that seems to add a more realistic and instructional (if constructed well) element to the previously monotonous selected-response instruments of the past. The problems that stem from assessments are not often due to the assessments themselves, but rather to the handling of the results. This is why it is still so critical to continue to use other measures in the evaluation of students, schools, districts, and states.

Having discussed a few issues and concerns regarding standards and assessments, we must look at the impact these efforts have had on instructional improvement. There has been some research and much discussion on whether this emphasis on standards for student performance and assessment systems has proven to be a positive venture. Comments span both

ends of a continuum. Some believe that there is so much emphasis placed on teaching to a test that overall student achievement has suffered because of the narrowed curriculum. There have been problems not only with being too item-focused or skill-specific, but also with being focused only on the students that can pull up the test results and relieve any fear of retribution for low scores.

Others, however, contend that, in certain circumstances, there is improved student achievement as a result of improved instruction. "New accountability systems that are well-designed (with fair, comprehensible, meaningful, and stable features) are associated with improved student achievement when adequate capacity to improve instruction is present in schools or can be provided by an outside partner" (Fuhrman, 1999, p. 10). The qualifiers in the parenthetical statement are not to be taken lightly. These are difficult elements to assure in a program. Fuhrman goes on to say, "in the absence of explicit attention to capacity, the new systems are insufficient approaches to improving student achievement" (1999, p. 10). The critical nature of professional development to provide teachers with the capacity to teach the standards proficiently needs to be at the forefront of most discussions and yet, often is ignored. This ability of the teachers to impart knowledge and provide assistance to those needing additional help, for the most part, will determine the success or failure of a "standards and assessment" program designed to increase student achievement. Again, it appears that, in many instances, much more time, effort, and funding are spent on developing and documenting the standards, than on providing the necessary assistance for the teaching of the identified essential skills.

In discussing "lessons from last decade's reforms," it has been noted that states with the highest test scores "have long supported high-quality teaching and teacher learning." These states do not necessarily have strict statewide curriculum or high-stakes testing programs, but they "do have a long history of professional policy. Reform strategies that did not make substantial efforts to improve teaching have been much less successful" (Darling-Hammond and Ball, 1998, p. 3). Teachers in the classroom must have the opportunity to secure the skills necessary for success. Therefore, teacher-training programs must become involved more adequately so that standards for students and standards for teachers are precisely aligned.

The underlying question is whether the emphasis on standards and assessment is having a positive effect on instructional improvement and overall achievement. Many say the verdict is still out, especially since there are so many places having to take to heart the effects of initial failure of some of the newer, tougher standards and the assessments that measure them. A few places are retreating from their initial requirements in light of all the students that will be impacted negatively, while others are crying they should not kill the messenger but stand by the rigorous advances.

So, are all of these efforts working? Yes and no. (Hasn't this always been true?) Some who have researched specific locations declare that these systems have "helped channel teachers' work to the most important goals of the system ..." and that some of the consequences have helped to motivate teachers to work in "more focused ways to produce improved student achievement" (Fuhrman, 1999, p. 6). Organizations such as Achieve are assisting in the dialogue

of helping states improve their systems and have been encouraged by the direction and commitments of some states and districts as well as their policymakers.

One significant problem in answering this question of effectiveness of the systems, lies in how often programs are changed before they can be declared successful or unsuccessful. Newly elected or appointed officials often want to make their presence known by restructuring or hastily changing the programs in place. We seem to start over constantly with something new before there has been sufficient effort and research to make an educated statement about the impact of a specific program. It may take three years to put a program in place, but positive results are expected within six months or the program is doomed. In addition, it is not uncommon for a new system to be mandated, yet, not adequately funded. Policymakers and the public, however, still hold the schools and districts accountable for the program's success.

Perhaps, the very best result of the efforts toward higher standards has been the dialogue created. When educators and policymakers gather and critically discuss the reality of the nature of our schools and how to improve what happens in classrooms, some good comes of it. Thousands of groups of teachers have been brought together to review and analyze their curriculum and discuss it across grade levels. This is not new, but it must continue if teachers are to keep abreast of what remains essential in curriculum and process. Each time there is a new effort that forces people to evaluate critically the current system, many are effected in a positive way--if simply by a deeper knowledge and expression of what they are doing and why.

This critical evaluation of classroom goals occurs at many levels, but when it occurs with teachers it can affect their students directly. One example is from an experience with the development of a graduation exam and one of the necessary tasks of the process to assure instructional and curricular validity. A statewide survey was conducted with teachers of various grade levels who were to answer questions regarding specific skills that they taught in their classrooms. The subject area was math and this particular incident was with seventh- and eighth-grade teachers. They were to answer whether they taught certain specific skills to their students. If the answer was no, they were to state why (the skill is too easy, the skill is too difficult for that grade level, etc.). The results for several skills astounded the teachers. Teachers had confirmed already that these skills were, indeed, essential skills for the curriculum. The seventh-grade teachers stated that there were several skills that they did not teach because they were too difficult for the grade level. The eighth-grade teachers declared that they did not teach these very same skills because they were too easy for the grade level! If nothing else, they certainly learned about the lack of communication and set out to establish exactly where these "essential skills" had fallen through the crack. Dialogue and critical evaluation by discussion create knowledge and understanding. An effective standards and assessment system generates this dialogue.

There are other benefits of these current efforts. There seems to be an attempt in many places to do a better job in bridging the disciplines and there is more emphasis, at least, on addressing the critical need for meaningful professional development. Another benefit has resulted from the fact that we must address more explicitly the issues of diverse populations in our schools. The cry for higher standards has brought more national attention to the need to provide for all students.

Still, there are many places where programs have not made an impact and, in some cases, may have created problems. There is still too much attention on a single score on a single test at the expense of overall achievement of all students. When these assessments dictate, the interpretation still seems to be extremely grade-oriented. If it is a sixth-grade test then the burden is on the sixth-grade teacher. This is astounding since we have been battling this since the days of minimum competency testing. An entire school must accept responsibility for the growth of a child and not lay the burden on only the grade levels tested. Something about "it takes a village ...". Funding efforts are often expended for the development of standards and especially for the assessments that measure them, ignoring the professional development that is needed to assure that the teachers have the capacity to teach well. Changes in subject areas, particularly with the onslaught of technology, demand that teachers have every opportunity to hone their skills and enhance their practices.

Identifying standards does not make "kids smarter." Administering assessments does not make "kids smarter." However, if a system is implemented with the goal of a more effective instructional program to maximize the potential of each student then a significant impact can occur. What happens in the classrooms is still the heart of the matter. One thing seems clear. Even those of us who feel there is a lot of "déjà vu all over again" realize, if we never embraced new efforts, discussions would cease, evaluation of current practice would diminish, and stagnation surely would result. We must keep everyone talking and analyzing if our programs are going to be successful. Are these programs working? In some places, yes in some places, no--just as in the past. Are these programs worthwhile? Absolutely. They keep us trying to find the answers.

References

- Darling-Hammond, L. & Ball, D. L. (1998). *What can policymakers do to support teaching to high standards?* (CPRE Bulletin) Philadelphia, PA: University of Pennsylvania, Consortium for Policy Research in Education.
- Fuhrman, S. H. (1999). *The new accountability*. (CPRE Policy Briefs No. RB-27). Philadelphia, PA: University of Pennsylvania, Consortium for Policy Research in Education.
- Gandal, M. (1996). *Making standards Matter, 1996: An Annual Fifty State Report on Efforts to Raise Academic Standards*. Washington, DC: American Federation of Teachers.
- Hoff, D. J. (2000, January 26). Testing ups and downs predictable. *Education Week*.
- O'Neil, J. (1999). Core knowledge & standards - A conversation with E. D. Hirsch. *Educational Leadership*, 56(6), 28-31.
- Popham, W. J. (1999). Where large scale educational assessment is heading and why it shouldn't. *Educational Measurement: Issues and Practices*, 18(3), 13-17.
- Popham, W. J. (1999). Why standardized tests don't measure educational quality. *Educational Leadership*, 56(6), 8-15.
- Schomaker, M. & Marzano, R. J. (1999). Realizing the promise of standards-based education. *Educational Leadership*, 56(6), 17-21.